

## PATENT COOPERATION TREATY

## From the INTERNATIONAL BUREAU

PCT

**NOTIFICATION OF ELECTION**

Assistant Commissioner for Patents  
United States Patent and Trademark  
Office  
Box PCT  
Washington, D.C.20231  
ETATS-UNIS D'AMERIQUE

<b>Date of mailing (day/month/year)</b> 15 March 2000 (15.03.00)	in its capacity as elected Office
<b>International application No.</b> PCT/CA99/00637	<b>Applicant's or agent's file reference</b> 80021-100
<b>International filing date (day/month/year)</b> 14 July 1999 (14.07.99)	<b>Priority date (day/month/year)</b> 14 July 1998 (14.07.98)
<b>Applicant</b>	SMIT, John

1. The designated Office is hereby notified of its election made:

in the demand filed with the International Preliminary Examining Authority on:

07 February 2000 (07.02.00)

in a notice effecting later election filed with the International Bureau on:

2. The election  was

was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

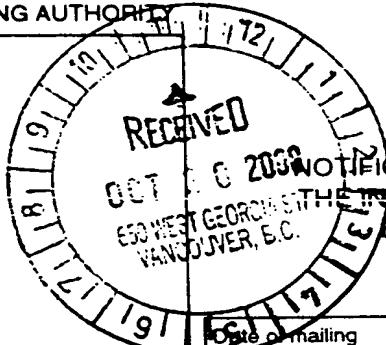
<p><b>The International Bureau of WIPO</b>  <b>34, chemin des Colombettes</b>  <b>1211 Geneva 20, Switzerland</b></p> <p>Facsimile No.: (41-22) 740.14.35</p>	<p><b>Authorized officer</b></p> <p><b>S. Mafla</b></p> <p>Telephone No.: (41-22) 338.83.38</p>
---	---

PATENT COOPERATION TREATY

From the  
INTERNATIONAL PRELIMINARY EXAMINING AUTHORITY

To:

ROBINSON, J. Christopher et al.  
SMART & BIGGAR  
Vancouver Centre, Suite 2200  
650 W. Georgia Street  
P.O. Box 11560  
Vancouver, BC V6B 4N8  
CANADA



PCT

Date of mailing  
(day/month/year)

10.10.2000

Applicant's or agent's file reference  
82104-2

**IMPORTANT NOTIFICATION**

International application No. PCT/CA99/00637	International filing date (day/month/year) 14/07/1999	Priority date (day/month/year) 14/07/1998
---	--	--

Applicant  
THE UNIVERSITY OF BRITISH COLUMBIA et al.

1. The applicant is hereby notified that this International Preliminary Examining Authority transmits herewith the international preliminary examination report and its annexes, if any, established on the international application.
2. A copy of the report and its annexes, if any, is being transmitted to the International Bureau for communication to all the elected Offices.
3. Where required by any of the elected Offices, the International Bureau will prepare an English translation of the report (but not of any annexes) and will transmit such translation to those Offices.

**4. REMINDER**

The applicant must enter the national phase before each elected Office by performing certain acts (filing translations and paying national fees) within 30 months from the priority date (or later in some Offices) (Article 39(1)) (see also the reminder sent by the International Bureau with Form PCT/IB/301).

Where a translation of the international application must be furnished to an elected Office, that translation must contain a translation of any annexes to the international preliminary examination report. It is the applicant's responsibility to prepare and furnish such translation directly to each elected Office concerned.

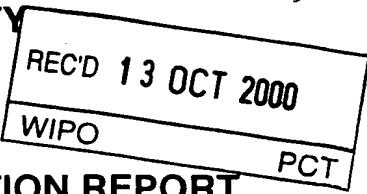
For further details on the applicable time limits and requirements of the elected Offices, see Volume II of the PCT Applicant's Guide.

*PLK*  
*LCR*

Name and mailing address of the IPEA/  European Patent Office D-80298 Munich Tel. +49 89 2399 - 0 Tx: 523656 eprmu d Fax: +49 89 2399 - 4465	Authorized officer  Vullo, C  Tel. +49 89 2399-8061	
---	---	---

## PATENT COOPERATION TREATY

PCT



## INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference 82104-2	<b>FOR FURTHER ACTION</b>		See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
International application No. PCT/CA99/00637	International filing date (day/month/year) 14/07/1999	Priority date (day/month/year) 14/07/1998	
International Patent Classification (IPC) or national classification and IPC C12N15/62			
<b>Applicant</b> THE UNIVERSITY OF BRITISH COLUMBIA et al.			

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.
2. This REPORT consists of a total of 5 sheets, including this cover sheet.

This report is also accompanied by ANNEXES, i.e. sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

These annexes consist of a total of sheets.

3. This report contains indications relating to the following items:

- I  Basis of the report
- II  Priority
- III  Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- IV  Lack of unity of invention
- V  Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- VI  Certain documents cited
- VII  Certain defects in the international application
- VIII  Certain observations on the international application

Date of submission of the demand 07/02/2000	Date of completion of this report 10.10.2000
Name and mailing address of the international preliminary examining authority: European Patent Office D-80298 Munich Tel. +49 89 2399 - 0 Tx: 523656 epmu d Fax: +49 89 2399 - 4465	Authorized officer Wimmer, G Telephone No. +49 89 2399 7347



**INTERNATIONAL PRELIMINARY  
EXAMINATION REPORT**

International application No. PCT/CA99/00637

**I. Basis of the report**

1. This report has been drawn on the basis of (substitute sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments.):

**Description, pages:**

1-26 as originally filed

**Claims, No.:**

1-8 as originally filed

2. The amendments have resulted in the cancellation of:

the description,      pages:  
 the claims,      Nos.:  
 the drawings,      sheets:

3.  This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

4. Additional observations, if necessary:

**V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**

**1. Statement**

Novelty (N)	Yes:      Claims 1-8
	No:      Claims
Inventive step (IS)	Yes:      Claims 1-8
	No:      Claims
Industrial applicability (IA)	Yes:      Claims 1-8
	No:      Claims

**INTERNATIONAL PRELIMINARY  
EXAMINATION REPORT**

International application No. PCT/CA99/00637

2. Citations and explanations

**see separate sheet**

**INTERNATIONAL PRELIMINARY  
EXAMINATION REPORT - SEPARATE SHEET**

International application No. PCT/CA99/00637

**Re Item V**

Reasoned statement under Rule 66.2(a)(ii) with regard to novelty, inventive step or industrial applicability.

**The application meets the requirements of Art. 33 PCT.**

- 1) Reference is made to the following documents (the document numbering corresponds to their order of citation in the international search report):

D1: WO 97 34000 A (UNIV BRITISH COLUMBIA ;SMIT JOHN (CA); BINGLE WADE H (CA); NOMELLI) 18 September 1997 (1997-09-18) cited in the application

D2: US-A-5 500 353 (SMIT JOHN ET AL) 19 March 1996 (1996-03-19)

D5: JOHN WILEY & SONS INC.: 'Introduction to expression by fusion protein vectors' CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, vol. suppl 28, no. chapter 16.4A, pages 16.4.1-16.16.4.17, XP002125707 cited in the application

D6: LANDON M. : 'Cleavage at Aspartyl-Prolyl bonds' METHODS IN ENZYMOLOGY, vol. XLVII, no. part E, 1977, pages 145-149, XP000863051 cited in the application

**Novelty under Art. 33(2) PCT.**

- 2) The methods of claims 7 and 8 refer to the creation of fusion proteins of a *Caulobacter* S-layer protein and a heterologous protein, where the fusion protein carries an Asp-Pro dipeptide.

The fusion proteins themselves are disclosed in documents D1 and D2. In these documents, the authors describe fusion proteins of a heterologous protein and the S-layer protein of *Caulobacter*. The DNA constructs for the creation of these fusion proteins contain linker sequences, including a BamHI restriction (GGATCC) site which, when translated in the proper reading frame, will lead to an Asp-Pro (GAT CCN) dipeptide. Since D1 and D2 describe such linkers in all three reading frames, fusion proteins with Asp-Pro dipeptides at the fusion peptide junction will

eventually be formed.

However, it was not the intention of the methods of D1 or D2 to create fusion proteins with Asp-Pro dipeptides, and neither of these documents points out the possibility of Asp-Pro dipeptide formation.

Novelty is therefore acknowledged for the methods of claims 7 and 8.

- 3) For similar reasons, and since documents D1 and D2 moreover do not describe the cleavage of the fusion proteins described therein by an acid solution, also claims 1-6 are novel.

Inventive Step under Art. 33(3) PCT.

- 4) The subject-matter of claim 1 extends to the method of cleaving fusion proteins of a protein of interest and the c-terminal secretion signal of the S-layer protein of Caulobacter, where the fusion protein contains an Asp-Pro dipeptide, and where the method comprises the application of low pH to cleave said Asp-Pro dipeptide.

As stated above, according fusion proteins were already disclosed by documents D1 and D2. However, as described above, this disclosure is regarded to be accidental. Moreover, the cleavage of a fusion protein comprising part of the Caulobacter S-layer protein is not mentioned in these documents.

Therefore, and although documents D5 and D6 describe the use of low pH as a standard method to cleave (fusion-)proteins containing an Asp-Pro dipeptide, the IPEA regards claim 1, and the dependent claims 2-6, as containing an inventive step.

- 5) Consequently, also claims 7 and 8, which are concerned with the provision of fusion proteins usable for the above method, are considered to contain an inventive step.

## PATENT COOPERATION TREATY

PCT

## INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference <b>80021-100</b>	<b>FOR FURTHER ACTION</b> see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.	
International application No. <b>PCT/CA 99/ 00637</b>	International filing date (day/month/year) <b>14/07/1999</b>	(Earliest) Priority Date (day/month/year) <b>14/07/1998</b>
Applicant <b>THE UNIVERSITY OF BRITISH COLUMBIA et al.</b>		

This International Search Report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This International Search Report consists of a total of 3 sheets.

It is also accompanied by a copy of each prior art document cited in this report.

## 1. Basis of the report

a. With regard to the **language**, the international search was carried out on the basis of the international application in the language in which it was filed, unless otherwise indicated under this item.

the international search was carried out on the basis of a translation of the international application furnished to this Authority (Rule 23.1(b)).

b. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international search was carried out on the basis of the sequence listing :

contained in the international application in written form.

filed together with the international application in computer readable form.

furnished subsequently to this Authority in written form.

furnished subsequently to this Authority in computer readable form.

the statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.

the statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished

**Certain claims were found unsearchable (See Box I).**

**Unity of Invention is lacking (see Box II).**

#### 4. With regard to the title

the text is approved as submitted by the applicant.  
 the text has been established by this Authority to read as follows:

5. With regard to the **abstract**,

the text is approved as submitted by the applicant.  
 the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. The figure of the drawings to be published with the abstract is Figure No.

- as suggested by the applicant.
- because the applicant failed to suggest a figure.
- because this figure better characterizes the invention.

None of the figures.

# INTERNATIONAL SEARCH REPORT

International Application No

PCT/CA 99/00637

**A. CLASSIFICATION OF SUBJECT MATTER**

IPC 7 C12N15/62 C12N15/74 C07K14/21 C07K14/00 C07K14/08  
C07K14/245

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 97 34000 A (UNIV BRITISH COLUMBIA ;SMIT JOHN (CA); BINGLE WADE H (CA); NOMELLI) 18 September 1997 (1997-09-18) cited in the application the whole document ----	1-8
Y	US 5 500 353 A (SMIT JOHN ET AL) 19 March 1996 (1996-03-19) the whole document ----	1-8 -/-

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

\* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

17 December 1999

Date of mailing of the international search report

11/01/2000

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Chambonnet, F

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/CA 99/00637

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	BINGLE W H ET AL: "LINKER MUTAGENESIS OF THE CAULOBACTER CRESCENTUS S-LAYER PROTEIN: TOWARD A DEFINITION OF AN N-TERMINAL ANCHORING REGION AND A C-TERMINAL SECRETION SIGNAL AND THE POTENTIAL FOR HETEROLOGOUS PROTEIN SECRETION" JOURNAL OF BACTERIOLOGY, US, WASHINGTON, DC, vol. 179, no. 3, February 1997 (1997-02), page 601-611 XP002034945 ISSN: 0021-9193 cited in the application page 604, column 1, line 5 -page 608, column 1, paragraph 2; figure 2 ----	1-8
Y	NOMELLINI J F ET AL: "INSERTION OF HETEROLOGOUS PEPTIDES WITHIN THE SURFACE-LAYER PROTEIN OF CAULOBACTER CRESCENTUS" ABSTRACTS OF THE GENERAL MEETING OF THE AMERICAN SOCIETY FOR MICROBIOLOGY, US, THE SOCIETY, WASHINGTON, DC, vol. 95, 1995, page 525 XP002035105 ISSN: 1060-2011 abstract ----	1-8
Y	JOHN WILEY & SONS INC.: "Introduction to expression by fusion protein vectors" CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, vol. suppl 28, no. chapter 16.4A, pages 16.4.1-16.4.17, XP002125707 cited in the application page 16.4.1 -page 16.4.4 page 16.4.13 ----	1-8
Y	LANDON M. : "Cleavage at Aspartyl-Prolyl bonds" METHODS IN ENZYMOLOGY, vol. XLVII, no. part E, 1977, pages 145-149, XP000863051 cited in the application the whole document -----	1-8

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International Application No

PCT/CA 99/00637

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO 9734000	A 18-09-1997	US 5976864	A	02-11-1999
		US 5500353	A	19-03-1996
		AU 2019497	A	01-10-1997
		CA 2175549	A	13-09-1997
		CA 2247805	A	18-09-1997
		EP 0888454	A	07-01-1999
		CA 2090549	A	10-12-1999
-----	-----	-----	-----	-----
US 5500353	A 19-03-1996	WO 9734000	A	18-09-1997
		US 5976864	A	02-11-1999
		CA 2090549	A	10-12-1993
-----	-----	-----	-----	-----



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : <b>C12N 15/62, 15/74, C07K 14/21, 14/00, 14/08, 14/245</b>		A1	(11) International Publication Number: <b>WO 00/04170</b> (43) International Publication Date: 27 January 2000 (27.01.00)
(21) International Application Number: <b>PCT/CA99/00637</b>		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: 14 July 1999 (14.07.99)		(30) Priority Data: 2,237,704 14 July 1998 (14.07.98) CA	
(71) Applicant (for all designated States except US): THE UNIVERSITY OF BRITISH COLUMBIA [CA/CA]; IRC Room 331, 2194 Health Sciences Mall, Vancouver, British Columbia V6T 1Z3 (CA).		(72) Inventor; and (75) Inventor/Applicant (for US only): SMIT, John [US/CA]; 9960 Seacastle Drive, Richmond, British Columbia V7A 4R8 (CA).	
(74) Agents: ROBINSON, Christopher, J. et al.; Suite 2200, 650 West Georgia Street, P.O. Box 11560, Vancouver, British Columbia V6B 4N8 (CA).		Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	

(54) Title: CLEAVAGE OF CAULOBACTER PRODUCED RECOMBINANT FUSION PROTEINS

## (57) Abstract

This invention provides a method for cleaving target proteins from Caulobacter S-layer protein under mild acid conditions. A fusion protein secreted by Caulobacter which includes a target protein and a Caulobacter S-layer secretion signal may be cleaved at an aspartate-proline dipeptide without solubilizing the fusion protein. This method may be carried out while the fusion protein is in an insoluble aggregate which facilitates recovery of the protein. This invention also provides a method of preparing a DNA construct for expression of the fusion protein and a method of preparing the fusion protein.

***FOR THE PURPOSES OF INFORMATION ONLY***

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LJ	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## CLEAVAGE OF CAULOBACTER PRODUCED RECOMBINANT FUSION PROTEINS

### 5 FIELD OF INVENTION

This invention relates to the expression and secretion of recombinant fusion proteins from Caulobacter wherein a heterologous polypeptide is fused with all or part of the surface layer protein (S-layer protein) of the bacterium.

10

### BACKGROUND OF THE INVENTION

Many bacteria assemble layers composed of repetitive, regularly aligned, proteinaceous sub-units on the outer surface of the cell. These layers are essentially 15 two-dimensional paracrystalline arrays, and being the outer molecular layer of the organism, directly interface with the environment. In Caulobacter, the S-layer protein is synthesized by the cell in large quantities and the S-layer completely envelops the cell and thus appears to be a protective layer.

Caulobacter are natural inhabitants of most soil and freshwater environments 20 and may persist in waste water treatment systems and effluents. The bacteria alternate between a stalked cell that is attached to a surface, and an adhesive motile dispersal cell that searches to find a new surface upon which to stick and convert to a stalked cell. The bacteria attach tenaciously to nearly all surfaces and do so without producing the extracellular enzymes or polysaccharide "slimes" that are characteristic of most other 25 surface attached bacteria. Caulobacters have simple requirements for growth. The organism is ubiquitous in the environment and has been isolated from oligotrophic to mesotrophic situations. They are known for their ability to tolerate low nutrient level stresses, for example, low phosphate levels.

All of the freshwater Caulobacter that produce an S-layer are similar and have 30 S-layers that are substantially the same under electron microscopy. The layers are hexagonally arranged in all cases, with a similar centre-centre dimension (see: Walker, S.G., et al... (1992). "Isolation and Comparison of the Paracrystalline Surface Layer Proteins of Freshwater Caulobacters" J. Bacteriol. 174: 1783-1792).

16S rRNA sequence analysis of several S-layer producing Caulobacter strains show that they group closely (see: Stahl, D.A. *et al.* (1992) "The Phylogeny of Marine and Freshwater Caulobacters Reflects Their Habitat" *J. Bacteriol.* 174: 2193-2198). DNA probing of Southern blots using the S-layer gene from C. crescentus CB15 5 identifies a single band that is consistent with the presence of a cognate gene (see: MacRae, J.D. and, J. Smit. (1991) "Characterization of Caulobacters Isolated from Wastewater Treatment Systems" *Applied and Environmental Microbiology* 57:751-758). Furthermore, antisera raised against the S-layer protein of CB15 reacts against the S-layer protein of other Caulobacter (see: Walker, S.G. *et al.* (1992) [*supra*]). All 10 S-layer proteins isolated from Caulobacter may be substantially purified using the same methods. All strains appear to have a polysaccharide species which may be required for S-layer attachment (see: Walker, S.G. *et al.* (1992) [*supra*]).

The S-layers elaborated by freshwater isolates of Caulobacter are visibly indistinguishable from the S-layer produced by Caulobacter strains CB2 and CB15. 15 The S-layer proteins from the latter strains have approximately 100,000 m.w. although sizes of S-layer proteins from other species and strains will vary. The hydrophilic S-layer protein has been characterized both structurally and chemically. It is composed of ring-like structures spaced at 22 nm intervals arranged in a hexagonal manner on the outer membrane. The S-layer is bound to the bacterial surface and may be removed by 20 low pH treatment or by treatment with a calcium chelator such as EDTA.

The similarity of S-layer proteins in different strains of Caulobacter permits the use of a cloned S-layer protein gene of one Caulobacter strain for retrieval of the corresponding gene in other Caulobacter strains (see: Walker, S.G. *et al.* (1992) [*supra*]; and MacRae, J.D. *et al.* (1991) [*supra*]).

25 Expression of a heterologous polypeptide as a fusion product with the S-layer protein of Caulobacter provides advantages not previously seen in systems for production of recombinant fusion proteins using other organisms such as E. coli and Salmonella. All known Caulobacter strains are believed to be harmless and are nearly ubiquitous in aquatic environments. In contrast, many Salmonella and E. coli strains 30 are pathogens. Consequently, expression and secretion of a heterologous polypeptide using Caulobacter as a vehicle has the advantage that the expression system will be

stable in a variety of outdoor environments and may not present problems associated with the use of a pathogenic organism. Furthermore, Caulobacter are natural biofilm forming species and may be adapted for use in fixed biofilm bioreactors. The quantity of S-layer protein that is synthesized and is secreted by Caulobacter is high, reaching 12%  
5 of the cell protein.

There is an existing need to produce pure proteins and peptides in an economical manner and in a manner that minimizes or simplifies the purification steps needed after fermentation. Key commercial areas include the production of recombinant human and animal therapeutic antibiotic and vaccine peptides, industrial enzymes, protein  
10 polymers, and antibacterial enzymes for foodstuffs. Many of these commercial applications require low production costs and there are few expression systems available that can meet such cost restraints. In addition, there are numerous research applications where rapid methods to produce and purify proteins are needed to facilitate the discovery stage. This is especially true where there is a desire to express a large  
15 number of proteins with unknown function (from a collections of cloned cDNA's, for example) or a large number of variants of a single protein, (for example, resulting from site directed mutagenesis) in a search for variants with improved properties.

Generally, proteins must be secreted to be produced at low cost. The primary reason is the much reduced cost of purification of the target protein from cell material.  
20 However, even for secreted proteins, simple methods of separating the product from spent culture and cells are important for cost reduction and ease of use.

An international patent application published as WO 97/34000 on September 18, 1997 describes the expression and secretion of recombinant proteins from Caulobacter in which the recombinant protein is a fusion of all or part of Caulobacter S-layer protein  
25 with a heterologous protein of interest (also see: Bingle, W.H., *et al.* 1997<sup>1</sup> "Linker Mutagenesis of the Caulobacter us S-layer protein: Toward a Definition of an N-terminal Anchoring Region and a C-terminal Secretion Signal and the Potential for Heterologous Protein Secretion". *J. Bacteriol.* 179:601-611).

The Caulobacter S-layer secretion apparatus is in the category of "Type 1" secretion usually found in pathogenic bacteria and noted for its ability to secrete a wide variety of proteins including large and hydrophilic proteins. The Caulobacter protein

secretion system is particularly useful to secrete recombinant proteins.

The Caulobacter S-layer Type 1 secretion pathway requires only a C-terminal secretion signal, typically comprising about 200 amino acids at the end of the protein. The export mechanism is capable of tolerating a wide variety of foreign proteins.

5 Recombinant proteins may be conveniently produced as fusion proteins with the target protein being fused to the C-terminal secretion signal. Depending on the application, it may be desirable to remove the secretion signal following secretion. Not removing the secretion signal may be an approach suitable for many subunit vaccine applications, where the remaining S-layer protein serves as a carrier.

10 A unique and desirable feature of fusion proteins produced by the Caulobacter S-layer protein secretion system is that they form insoluble aggregates in the culture medium. This is apparently a consequence of the S-layer sequences associated with secretion signal and reflects the fact that the protein normally self-assembles into a two dimensional crystalline layer on the bacterium's surface. These aggregates are visible 15 to the naked eye and are readily collected by simple filtration. With simple water wash steps, residual bacterial cells are readily flushed away. It is routinely possible to achieve a protein purity of 90% or better with this simple purification procedure.

#### DESCRIPTION OF THE PRIOR ART

20

Most current protein purification systems for recombinant proteins produced by bacteria rely upon an affinity matrix to achieve separation of the target protein and to concentrate the protein for subsequent steps of purification. To accomplish this, genes 25 for recombinant proteins are commonly constructed so that they contain affinity tags, which are protein sequences that will bind to an affinity matrix. Commonly used systems include the following:

(a) glutathione S-transferase (GST) tag, which binds to glutathione-sepharose matrices;

30

(b) maltose binding protein (MBP) tag, which binds to amylose matrices;

- (c) multiple tandem histidine residues (e.g. "His-6") tag, which binds to nickel-derivatized solid matrices; and
- 5 (d) protein A tag, which binds to Immunoglobulin IgG-derivatized sepharose or comparable matrices.

Prior art techniques were typically developed so that removal of a target protein does not disrupt the tag and matrix association. Instead, enzymes that cleave specific 10 sequences of amino acids are employed. The enzyme cleavage sequence is positioned between the tag and the desired recombinant protein and enzymatic cleavage is effected directly on the matrix with attached fusion protein. If a secretion signal is used, the cleavage site is usually positioned such that the secretion signal is separated from the target recombinant protein during the cleavage step. The matrix is regenerated for re- 15 use only after the target recombinant protein has been purified away from the matrix. Typical enzymes used in these methods are Factor Xa, enterokinase and collagenase.

Chemical cleavage is generally not used because the conditions required for 20 cleavage will disrupt the binding of affinity tag and matrix or destroy the matrix. When chemical cleavage is used with recombinant fusion proteins to cleave target protein from a secretion signal and/or affinity tag, solubilization and denaturation processes are generally employed. The expectation is that complete or nearly complete unfolding of the protein is a prerequisite for effective cleavage.

Mild-acid cleavage is predicated on the inclusion, by happenstance or design, of 25 the acid-sensitive aspartate-proline dipeptide at a desired site for cleavage. The protein to be cleaved is typically exposed to conditions that solubilize and/or completely denature the protein prior to cleavage. The chaotropic agent guanidine hydrochloride (used at 6-7 M) is commonly employed to denature and solubilize the protein prior to, or at the same time as acid treatment. Alternately, high concentrations of acids that also serve as solubilizing agents (as examples: 70-90% formic acid, acetic acid [10%] 30 pyridine, or relatively high concentrations of HCL (60 mM or more) are employed. Because such conditions would disrupt a tag/affinity matrix association, direct cleavage

of an affinity tag from the target protein while a protein remains associated with an affinity matrix is not attempted.

General conditions for cleavage at aspartate - proline sites are described in

5 Current Protocols in Molecular Biology (supp. 28; chapter 16.4) John Wiley & Sons Inc. 1994, and in Landon, M. "Cleavage at Aspartyl - Prolyl Bonds" in Methods in Enzymology (1977) 47: 145-149. These references suggest that significant variability of cleavage conditions exist for different proteins and that cleavage might occur in some instances without first denaturing or solubilizing the protein. However, in practice, the

10 latter circumstances are rare and proteins to be subjected to acid cleavage at Asp-Pro dipeptides are usually solubilized to a state where there is no visible turbidity. Such solubilized protein will normally not pellet when centrifuged at 100,000 x g for 1 hour. It is now shown that mild-acid conditions may be used for cleavage of aspartate-proline sites in Caulobacter S-layer fusion proteins without placing the protein in a solubilized

15 state as described above.

## SUMMARY OF INVENTION

This invention is based on the unexpected discovery that recombinant fusion

20 proteins produced by the Caulobacter S-layer protein secretion system can be cleaved under mild-acid conditions and solubilization of the fusion protein is not required. Cleavage may be accomplished while the fusion protein is in the form of an insoluble aggregate typical of the Caulobacter S-layer protein. Cleavage occurs at aspartate-protein dipeptides which may be in a heterologous protein portion of the fusion protein

25 or in a portion that is native to the Caulobacter S-layer portion. The dipeptide may be placed at a desired location for cleavage by engineering DNA encoding the fusion protein to express the dipeptide at the desired location. A preferable location for cleavage may be at or near the junction between a heterologous (target) protein and the Caulobacter S-layer portion comprising the Caulobacter secretion signal, such that a

30 cleavage product will be the target protein in its entirety and substantially free of extraneous amino acids.

The current invention makes it possible to cleave a heterologous (target) protein from the S-layer protein portion using only mild-acid conditions, even while the fusion protein is in an aggregated form. These cleavage conditions do not result in significant solubilization of the S-layer protein portion.

5

This invention provides a method of cleaving a fusion protein including a first component which comprises all or part of a Caulobacter S-layer protein including a Caulobacter C-terminal secretion signal, and a second component heterologous to Caulobacter. The fusion protein contains at least one aspartate-proline dipeptide. The 10 method comprises combining the fusion protein with an acid solution of a strength insufficient to solubilize the fusion protein for a time sufficient for cleavage of the fusion protein at the aspartate-proline dipeptide. The acid solution may have a pH of from about 1.5 (eg.  $1.5 \pm 0.1$ ) to about 2.5 (eg.  $2.5 \pm 0.1$ ), and preferably from about 1.65 (eg.  $1.65 \pm 0.05$ ) to about 2.35 (eg.  $2.35 \pm 0.05$ ). Preferred pH conditions may 15 be achieved using an acid equivalent in the range of about 5 to about 20 mM HCL. The method is typically carried out at a temperature in the range of approximately room temperature to about 50°C.

This invention also provides a method of preparing a DNA construct suitable for expression of a fusion protein suitable for use in the method of this invention. The 20 method comprises joining an upstream DNA segment including DNA heterologous to Caulobacter which includes a protein of interest to a downstream DNA segment including DNA for a Caulobacter C-terminal secretion signal which does not encode an aspartate-proline dipeptide. The upstream segment contains DNA encoding an aspartate-proline dipeptide at or near the junction between said upstream and 25 downstream segments.

This invention also provides a method of preparing a fusion protein, comprising the steps of expressing a DNA construct as described above in Caulobacter and recovering said fusion protein once secreted by the Caulobacter.

Once cleavage is accomplished according to this invention, the S-layer portion 30 comprising the Caulobacter secretion signal may remain as an insoluble aggregate. If the target protein is soluble, the S-layer portion may be easily separated from the target

recombinant protein by simple centrifugation or filtration methods. Thus the system of this invention facilitates separation as would a Tag/affinity matrix system except that here, the system is also the means for producing an insoluble matrix. In addition, the insoluble matrix produced by this invention is resistant to the effects of the acid 5 treatment, allowing direct cleavage of the target recombinant protein. In this way, a very inexpensive chemical cleavage method can be employed to economically retrieve recombinant proteins from a bacterial fusion protein. In contrast to the cost of most affinity matrices, there is little expense associated with the use of the S-layer secretion signal as it is simply a part of the fermentation/secretion process.

10

## DETAILED DESCRIPTION OF EMBODIMENTS OF THE INVENTION

### Production of Recombinant Fusion Proteins Using the *Caulobacter* S-layer Secretion System

15

Proteins may be produced using the *Caulobacter* S-layer Type 1 secretion pathway which requires only the C-terminal secretion signal of the *Caulobacter*. This signal is the C-terminal portion of the S-layer protein, which typically comprises about 200 amino acids. (See: Bingle, et al. (1997) [*supra*]; and, WO 97/34000). Additional 20 *Caulobacter* S-layer DNA upstream from the secretion signal may also be present and may be desirable to encode portions of the S-layer protein which will contribute to aggregate formation of the secreted protein. Such additional *Caulobacter* DNA may constitute most or all of the remainder of the DNA encoding the S-layer protein.

Standard techniques (such as methods described in WO 97/34000) may be used 25 to identify the amount of the C-terminal portion of a particular *Caulobacter* S-layer protein which functions as the secretion signal.

Creation of fusion proteins is commonly done by preparing DNA which codes for the target protein and fusing it in-frame with the C-terminal region of the S-layer gene. There are numerous possible methods, with the following being examples. 30 1. **Oligonucleotide Chemical Synthesis.** This involves the design of complementary single strands, complete with desirable restriction endonuclease cut sites

at the ends, chemical synthesis of the strands followed by annealing, cloning into a plasmid vector, juxtaposed to an appropriate portion of the C-terminal region of the S-layer gene.

**2. Production of the Target Gene DNA by Polymerase Chain Reaction (PCR)**

5 **Amplification of a Target Sequence.** In this case, appropriate in-frame restriction sites are incorporated into the short oligonucleotides used for amplification of a target sequence, such that the final PCR product can be treated with the appropriate restriction enzymes (to create the restriction site "sticky ends"), followed by cloning into a plasmid vector, juxtaposed to an appropriate portion of the C-terminal region of the S-layer 10 gene.

**3. Adapting Restriction Endonuclease Cleavage Sites that are Native to a Target Protein Gene Sequence for Fusion to the DNA Coding for the C-terminal S-layer Secretion Signal to Accomplish In-frame Expression of a Chimeric Protein.**

15 This can be accomplished by direct ligation (although it is uncommon that an appropriate match will occur), or the use of adapter sequences or methods involving blunting of a restriction site and subsequent blunt-end ligation to change expression reading frame or join unlike restriction site sticky ends.

There will be numerous convenient sites for fusion with the C-terminal regions 20 of the S-layer that lead to the successful expression, secretion and aggregation of a recombinant fusion protein. Some example positions are at or near the DNA sites corresponding to amino acids 622, 690, 784, 892 and 907 of the C. crescentus S-layer gene (see: Appendix 1 and, WO 97/34000). Other sites of fusion with the S-layer gene 25 may also be employed. Most often a plasmid vector is designed such that the C-terminal gene segment is resident on a plasmid with appropriate restriction sites placed at the N-terminal junction of the S-layer fragment. Target recombinant protein gene segments are then cloned into those restriction sites. It is typical to prepare initial plasmid constructs that are replicated in E.coli. After a construct is produced, it is typically transferred to a broad host range plasmid which can then be introduced into 30 the appropriate Caulobacter strain by electroporation. Suitable broad host range plasmids can be constructed from (but are not limited to) the IncQ, IncW and IncP1

plasmid incompatibility groups.

The introduction of the aspartate-proline (Asp-Pro) dipeptide at the appropriate site in the fusion protein can be done in several ways. Some examples are:

5 (a) incorporating a DNA sequence necessary to express the Asp-Pro dipeptide into the oligonucleotides used to prepare the target sequence, either by oligonucleotide synthesis or PCR methods;

10 (b) preparing a DNA segment with appropriate restriction sites at the termini so that an Asp-Pro dipeptide can be introduced (most often at the junction between S-layer and target gene) after a fusion recombinant S-layer gene has been made; and

15 (c) use of a native Asp-Pro dipeptide in either the target DNA or the S-layer segment (for example, an Asp-Pro dipeptide is located at amino acids 692 and 693 of the C. crescentus S-layer gene and is suitable for fusions made at the amino acid site).

20 The methods described above are not the only methods that may be used for creating and expressing fusion recombinant S-layer proteins, nor is it necessary to have the engineered genes resident on a plasmid. For example, the expressed gene may be introduced into the chromosome (using well-known gene insertion or replacement techniques) and still achieve secretion of the recombinant proteins (see WO 97/34000).

25 In some cases it may be desirable to produce recombinant fusion proteins as insertions of heterologous DNA in the middle of the S-layer gene. In such a case, Asp-Pro dipeptide sequences could be engineered at the N and C-termini of the target peptide.

All possible codon combinations for Asp-Pro will work but the CCA codon for proline is not preferred due to the likelihood of a low amount of the corresponding tRNA being present in Caulobacter. The following is an approximate usage table for C. crescentus.

TABLE 1

5

**Caulobacter crescentus Codon Usage Table**  
 [Amino Acid] [Triplet Code] [Frequency Per Thousand]

10	Phe UUU	2.5	Ser UCU	1.2	Try UAU	6.6	Cys UGU	0.6
	Phe UUC	27.0	Ser UCC	8.5	Try UAC	9.6	Cys UGC	5.5
	Leu UUA	0.0	Ser CA	1.2	STOP UAA	0.8	Cys UGA	1.6
	Leu UUG	4.4	Ser UCG	25.7	STOP UAG	0.6	STOP UGG	7.2
15	Leu CUU	4.4	Pro CCU	2.5	His CAU	3.2	Arg CGU	7.6
	Leu CUC	15.7	Pro CCC	15.5	His CAC	12.2	Arg CGC	44.7
	Leu CUA	1.1	Pro CCA	0.9	Gln CAA	3.7	Arg CGA	3.0
	Leu CUG	72.3	Pro CCG	27.1	Gln CAG	30.2	Arg CGG	12.1
20	IleAUU	2.4	Thr ACU	1.2	Asn AAU	4.1	Ser AGU	0.8
	Ile AUC	49.0	Thr ACC	37.3	Asn AAC	23.8	Ser AGC	14.9
	Ile AUA	0.3	Thr ACA	0.8	Lys AAA	2.7	Arg AGA	0.4
	Met AUG	25.7	Thr ACG	16.8	Lys AAG	37.9	Arg AGG	1.1
25	Val GUU	5.4	Ala GCU	9.5	Asp GAU	11.1	Gly GGU	9.5
	Val GUC	42.7	Ala GCC	84.1	Asp GAC	48.5	Gly GGC	64.8
	Val GUA	1.0	Ala GCA	2.2	Glu GAA	20.5	Gly GGA	2.3
	Val GUG	30.7	Ala GCG	36.7	Glu GAG	45.4	Gly GGG	7.7

Large quantities (eg. 12% of total cell protein/3% of input organic carbon) of a wide range of proteins can be produced, with yields in the order of 250 mg/liter of batch culture. Fusion proteins with 35 kDa of target peptide are secreted with little 5 difficulty, although proteins with multiple cysteines may be more difficult to express. Post-expression glycosylation of proteins does not occur, an advantage for most peptide expression applications.

#### 10 Host Expression Strains

For secretion of recombinant fusion S-layer proteins, the Caulobacter strain will preferably be one which has lost the ability to produce a native S-layer protein, while retaining a fully functional S-layer protein secretion apparatus. Such strains may be 15 obtained by screening for mutants that have spontaneously become S-layer protein negative; or, by directed genetic manipulation, such as (but not limited to) the insertion of a drug resistance cassette in the middle of the S-layer gene or the substitution of a version of the S-layer gene which has had a sizeable internal region deleted from the gene (see: Bingle *et al.* 1997<sup>1</sup> [*supra*]; Bingle *et al.* 1997<sup>2</sup> "Cell Surface Display of a 20 Pseudomonas aeruginosa PAK Pilin Peptide with the Paracrystalline Layer of Caulobacter crescentus" *Molec. Microbiol.* 26:277-288; and, Edwards and Smit (1991) "A Transducing Bacteriophage for Caulobacter us Uses the Paracrystalline Surface Layer Protein as a Receptor" *J. Bacteriol.* 173: 5568-5572). In the case of a genetic manipulation, a common method for producing such strains is to modify a copy of the 25 S-layer gene while on a plasmid and then to use well known gene replacement methods to substitute the modified gene for the native gene in the Caulobacter chromosome (see: Edwards and Smit (1991) [*supra*]).

If an entire S-layer gene is to be used for production of a recombinant protein (via insertion of a target sequence), strains defective in the production of the 30 lipopolysaccharide (LPS) used for S-layer attachment to the bacterial surface can be used. These can be prepared by forcing Caulobacter to grow without exogenous

calcium. Under these conditions mutants arise that are uniformly defective in producing a proficient version of the S-layer LPS (see: Walker, S.G. *et al.* (1994) "Characteristics of Mutants of Caulobacter crescentus Defective in Surface Attachment of the Paracrystalline Layer" *J. Bacteriol.* 176: 6312-6323).

5 All Caulobacter S-layer producing strains are suitable for this technology. One may isolate the S-layer gene from a particular strain (using homology between Caulobacter S-layers to design probes to detect and clone the S-layer genes) and adapt the C-terminal region for recombinant protein expression, in a manner similar to that done for C. crescentus strains (see: MacRae and Smit (1991) [*supra*], and Walker, S.G.  
10 *et al.* (1992) [*supra*]). Alternatively, one may construct recombinant fusion S-layer genes using the C. crescentus S-layer gene and express the recombinant genes in alternate Caulobacter hosts.

Freshwater Caulobacter producing S-layers may be readily detected by negative stain transmission electron microscopy techniques. Caulobacter may be isolated using  
15 the methods outlined by MacRae and Smit (1991) [*supra*], which take advantage of the fact that Caulobacter can tolerate periods of starvation while other soil and water bacteria may not and that they all produce a distinctive stalk structure, visible by light microscopy (using either phase contrast or standard dye staining methods). Once Caulobacter strains are isolated in a typical procedure, colonies may be suspended in  
20 2% ammonium molybdate negative stain and applied to plastic-filmed, carbon-stabilized 300 or 400 mesh copper or nickel grids and examined in a transmission electron microscope at 60 kilovolt accelerating voltage (see: Smit, J. (1986) "Protein Surface Layers of Bacteria", in Outer Membranes as Model Systems, (M. Inouge, ed. J. Wiley & Sons, at p. 343-376). S-layers are seen as two-dimensional geometric patterns most  
25 readily on those cells in a colony that have lysed and released their internal contents.

### Recombinant Protein Purification

Secreted proteins are separated and shed into the culture media as a macroscopic  
30 precipitate (the "aggregate" referred to herein). The shedding phenomenon is a consequence of the absence of the N-terminal region of the S-layer protein in the

expressed recombinant protein, or the loss of the lipopolysaccharide species used for S-layer attachment by the Caulobacter (see: Walker, S.G. et al. (1994) [*supra*]). Typically, the aggregate forms as loose, gel-like lumps of pure protein that can readily be retrieved and separated from the bacteria by simple filtration.

5 The aggregate may be readily separated from a soluble cleaved target protein by any suitable techniques such as filtration or centrifugation. If the target protein is insoluble once cleaved, it may then be convenient to then solubilize one or both of the proteins (for example in 8M urea or 6M quanidine HCL) and separate by chromatography. In this way, only 2 species of protein need to be separated.

10

### **Cleavage of Fusion Proteins**

General procedures for performing mild-acid cleavage are known from in the prior art as described above. In the method of this invention, conditions are adjusted to 15 avoid destruction of the target protein or solubilization of the aggregate containing the S-layer secretion signal. Excess acid or too high a temperature may increase the occurrence over time of random cleavages along the length of the fusion protein, which is to be avoided since such random cleavages may lead to undersized fragmentation of the fusion protein or solubilization of the aggregated S-layer portion.

20

Good yields of target protein with minimum random breaks in the fusion protein may generally be achieved by using from 5-20 mM HCL (or its equivalent while employing another acid). The respective pH of these conditions (unbuffered acid solution) is from about 2.3 to about 1.7. Time and temperature is preferably adjusted 25 by routine monitoring to achieve the desired cleavage while minimizing random breaks.

For example, temperature may range from room temperature to about 50° C. Time of treatment may range from about 12 to about 72 hours. Time or temperature outside of these ranges is permissible depending upon the strength of the acid and the accepted yield. Generally, lower yields are obtained with less acid strength, less time or lower 30 temperatures.

In the following examples, efficiency of cleavage in the order of 40-80% is

achieved using conditions the same as or similar to the following alternatives:

- 5 mM HCL at 50° C. for 48-72 hours
- 20 mM HCL at 30° C. for 48-72 hours.

Conditions in excess of the aforementioned values may be employed in some 5 cases with the possibility of random breaks increasing, particularly with increased acid strength or temperature. In the following examples, significant random cleavage occurred with 50 mM HCL at 50° C. after 48 hours.

Any acid may be employed in this invention which is normally used in solutions to which proteins are exposed. Acids which have a deleterious effect on proteins under 10 dilute conditions should be avoided. For example, HCL or an equivalent amount of H<sub>2</sub>SO<sub>4</sub> may be used in this invention but oxidizing acids such as nitric acid may not be suitable.

15 **Example 1. Cleavage of artificial silk protein sequences**  
**from a secretion signal containing a native aspartate-proline cleavage site.**

An artificial protein sequence resembling spider silk was constructed by synthesis of partially overlapping and complementing oligomers of DNA, which were then completed to a full duplex DNA with Taql polymerase extension, to create a 20 sequence that coded for 97 amino acids. The resulting DNA sequence and corresponding amino acid sequence are shown in Appendix 2.

The DNA sequence shown in Appendix 2 was cloned into a gene carrier sequence residing in a pUC8 plasmid cloning vector. The gene segment carrier had BamH1 restriction sites at each end and an internal BgIII site. This combination of 25 restriction sites allowed the production of multimers of the above sequence, relying on the fact that BamH1 sticky ends will ligate into BgIII sticky end, with the loss of both restriction sites. Thus one copy of the silk-like sequence within the gene segment carrier can be put inside a second copy of the same to produce a dimer. Using this principle, an 8X repeat was produced, fused to DNA encoding the S-layer secretion 30 signal corresponding to the C-terminal portion of the C. crescentus S-layer protein from about amino acid 690 onwards (see: Appendix 1). This fusion protein gene was

introduced into strain CB2A on a broad host range plasmid vector. The 8x multimer appeared to be unstable, resulting in recombination events that reduced the 8X multimer to a 3x size. The 3 fold repeat of the above 97 amino acid sequence, fused to the S-layer secretion signal was secreted. Protein was collected and subjected to treatment 5 with 5mM HCL for 2 days at 50° C. The result was the liberation of about 80% of soluble silk-like polymer which was readily separated by filtration from the S-layer protein which remained completely aggregated under these conditions. Cleavage occurred at native aspartate-proline dimer in the Caulobacter S-layer signal region (see: Appendix 1, amino acids numbered 692-693).

10

**Example 2. Cleavage of the salmonid virus Infectious Pancreatic Necrosis Virus (IPNV) surface glycoprotein candidate vaccine sequence from an S-layer secretion signal containing a native aspartate-proline site.**

15

The surface glycoprotein of the IPNV strain is a vaccine candidate. For this example and Example 4, the sequence of the first 257 amino acids of the mature protein and the corresponding DNA sequence as shown in Appendix 3 were used.

DNA encoding a segment of the major surface glycoprotein gene of IPNV 20 specifying amino acids 145-257 of the protein was fused to DNA sequence specifying two putative T-cell activating epitopes: MVF (SEQ ID No:1; LSEIKGVIVHRLEGV, derived from Measles Virus protein F) and P2 (SEQ ID No:2; QYIKANSKFIGITEL, derived from tetanus toxoid protein). The T-cell epitopes were positioned on the C-terminal end of the IPNV sequence. This chimeric protein was in turn fused in frame 25 with the C-crescentus S-layer gene at about amino acid 690 position of the gene and introduced into Caulobacter on a broad host range plasmid vector. The resulting secreted protein was collected and treated with 5 mM HCL for 2 days at 50° C. Cleavage occurred at the native aspartate-proline dimer described in Example 1. The result was the liberation of about 75% of soluble vaccine candidate chimeric protein 30 from the S-layer secretion signal which remained aggregated.

**Example 3. Cleavage of segments of an E. coli type I pilus tip subunit from an S-layer secretion signal containing a native aspartate-proline cleavage site.**

5        The FimH gene product is the tip pilus subunit of the E. coli strains involved with urinary tract infections. Two segments, T3 (specifying the first 145 amino acids of the mature peptide) and T7 (specifying the entire 258 amino acids of the mature peptide) were fused to the S-layer secretion signal at about amino acid 690 of the S-layer sequence. The T3 and T7 sequences are shown in Appendix 4.

10      The fusion protein genes were introduced into strain CB2A on a broad host range plasmid vector. In both cases the resulting secreted protein was collected and treated with 5 mM HCL for 2 days at 50° C. In both cases, the result was the liberation of about 50% of soluble vaccine candidate chimeric protein from the S-layer secretion signal which remained aggregated. Cleavage occurred at the native aspartate-proline 15 dimer described in Example 1.

**Example 4. Cleavage of the salmonid virus IPNV surface glycoprotein candidate vaccine sequence from an S-layer secretion signal containing an introduced aspartate-proline cleavage site.**

20      A segment of the major surface glycoprotein gene of IPNV specifying amino acids 1-257 of the protein shown in Appendix 4 was fused to a DNA sequence specifying a peptide containing an aspartate-proline dipeptide (SEQ ID No: 3; SPLGPAGDPEAS) such that the aspartate-proline dipeptide was positioned very near 25 the C-terminus of the chimeric protein. This chimeric protein was in turn fused in frame with the C. crescentus S-layer gene at about amino acid 784 position of the gene and introduced in strain CB2A on a broad host range plasmid vector. The resulting secreted protein was collected and treated with 5 mM HCL for 2 days at 50° C. Cleavage occurred at the introduced aspartate-proline dipeptide. The result was the 30 liberation of about 40% of insoluble vaccine candidate chimeric protein from the S-layer secretion signal which remained aggregated.

Longer DNA and amino acid sequences referred to above are set out in the

following Appendices which are part of this description. Appendix 1 sets out the complete nucleotide sequence of the C. crescentus S-layer gene (SEQ ID No: 4) with the upstream sequence including the -35 and -10 sites of the promoter region and the Shine Dalgarno sequence. The start codon is at nucleotide 101 and the coding sequence 5 run to and includes nucleotide 3179. The amino acid sequence of the C. crescentus S-layer protein (SEQ ID No: 5) included in Appendix 1 is predicted from the DNA sequence. Appendix 2 sets out the artificial spider silk DNA sequence (SEQ ID No:6) used in Example 1 and the corresponding amino acid sequence (SEQ ID No. 7). Appendix 3 sets out the DNA sequence (SEQ ID No: 8) and corresponding amino acid 10 sequence (SEQ ID No: 9) of the first 257 amino acids of IPNV as described in Examples 2 and 4. Appendix 4 sets out the T3 protein sequence (SEQ ID No: 10) and the T7 protein sequence (SEQ ID No: 11) as described in Example 3.

All publications, patents and patent applications referred to herein are hereby incorporated by reference. While this invention has been described according to 15 particular embodiments and by reference to certain examples, it will be apparent to those of skill in the art that variations and modifications of the invention as described herein fall within the spirit and scope of the attached claims.

19  
Appendix 1

GCTATTGTCG	ACGTATGACG	TTTGCTCTAT	AGCCATCGCT	GCTCCCATGC	GCGCCACTCG	60
GTGCGAGGGG	GTGTGGGATT	TTTTTTGGGA	GACAATCCTC	ATGGCCTATA	CGACGGCCCA	120
GTTGGTGACT	GCGTACACCA	ACGCCAACCT	CGGCAAGGCG	CCTGACGCCG	CCACCACGCT	180
GACGCTCGAC	GCGTACCGA	CTCAAACCCA	GACGGGCGGC	CTCTCGGACG	CCGCTGCGCT	240
GACCAACACC	CTGAAGCTGG	TCAACAGCAC	GACGGCTGTT	GCCATCCAGA	CCTACCAGTT	300
CTTCACCGGC	GTTGCCCCGT	CGGCCGCTGG	TCTGGACTTC	CTGGTCGACT	CGACCACCAA	360
CACCAACGAC	CTGAACGACG	CGTACTACTC	GAAGTTCGCT	CAGGAAAACC	GCTTCATCAA	420
CTTCTCGATC	AACCTGGCCA	CGGGCGCCGG	CGCCGGCGCG	ACGGCTTTCG	CCGCCGCCTA	480
CACGGGCGTT	TCGTACGCC	AGACGGTCGC	CACCGCTTAT	GACAAGATCA	TCGGCAACGC	540
CGTCGCGACC	GCCGCTGGCG	TCGACGTCGC	GGCCGCCGTG	GCTTTCCTGA	GCCGCCAGGC	600
CAACATCGAC	TACCTGACCG	CCTTCGTGGG	CGCCAACACG	CCGTTCACCG	CCGCTGCCGA	660
CATCGATCTG	GCCGTCAAGG	CCGCCCTGAT	CGGCACCATC	CTGAACGCCG	CCACGGTGTC	720
GGGCATCGGT	GGTTACCGA	CCGCCACGGC	CGCGATGATC	AACGACCTGT	CGGACGGCGC	780
CCTGTCGACC	GACAACGCGG	CTGGCGTGAA	CCTGTTCAAC	GCCTATCCGT	CGTCGGCGT	840
GTCGGGTTCG	ACCCCTCTCGC	TGACCACCGG	CACCGACACC	CTGACGGGCA	CCGCCAACAA	900
CGACACGTTT	GTTGCGGGTG	AAAGTCGCCGG	CGCTGCGACC	CTGACCGTTG	GCGACACCCCT	960
GAGCGGCGGT	GCTGGCACCG	ACGTCCCTGAA	CTGGGTGCAA	GCTGCTGCCG	TTACGGCTCT	1020
GCCGACCGGC	GTGACGATCT	CGGGCATCGA	AAAGATGAAC	GTGACGTCGG	GCGCTGCGAT	1080
CACCCCTGAAC	ACGTCTTCGG	CGGTGACGGG	TCTGACCGCC	CTGAACACCA	ACACCAGCGG	1140
CGCGGCTCAA	ACCGTCACCG	CGGGCGCTGG	CCAGAACCTG	ACCGCCACGA	CCGCCGCTCA	1200
AGCCGCGAAC	AACGTCGCCG	TCGACGGGCG	CGCCAACGTC	ACCGTCGCCT	CGACGGCGT	1260
GACCTCGGGC	ACGACCAACGG	TCGGCGCCAA	CTCGGCCGCT	TCGGGCACCG	TGTCGGTGAG	1320
CGTCGCGAAC	TCGAGCACGA	CCACCACGGG	CGCTATCGCC	GTGACCGGTG	GTACGGCCGT	1380
GACCGTGGCT	CAAACGGCCG	GCAACGCCGT	GAACACCACG	TTGACGCAAG	CCGACGTGAC	1440
CGTGACCGGT	AACTCCAGCA	CCACGGCCGT	GACGGTCACC	CAAACCGCCG	CCGCCACCGC	1500
CGGCGCTACG	GTCGCCGGTC	GCGTCAACCGG	CGCTGTGACG	ATCACCGACT	CTGCCGCCGC	1560
CTCGGCCACG	ACCGCCGGCA	AGATCGCCAC	GGTCACCCCTG	GGCAGCTTCG	GCGCCGCCAC	1620
GATCGACTCG	ACCGCTCTGA	CGACCGTCAA	CCTGCGGGC	ACGGGCACCT	CGCTCGGCAT	1680

20  
Appendix 1 (cont'd)

CGGCCGCGGC	GCTCTGACCG	CCACGCCGAC	CGCCAACACC	CTGACCCCTGA	ACGTCAATGG	1740
TCTGACGACG	ACCGGCGCGA	TCACGGACTC	GGAAGCGGCT	GCTGACGATG	GTTTCACCAC	1800
CATCAACATC	GCTGGTTCGA	CCGCCTCTTC	GACGATCGCC	AGCCTGGTGG	CCGCCGACGC	1860
GACGACCCCTG	AACATCTCGG	GCGACGCTCG	CGTCACGATC	ACCTCGCACA	CCGCTGCCGC	1920
CCTGACGGGC	ATCACGGTGA	CCAACAGCGT	TGGTGCACC	CTCGGCCCG	AACTGGCGAC	1980
CGGTCTGGTC	TTCACGGGCG	GGCCTGGCCG	TGACTCGATC	CTGCTGGCG	CCACGACCAA	2040
GGCGATCGTC	ATGGGCGCCG	GCGACGACAC	CGTCACCGTC	AGCTCGGCGA	CCCTGGCGC	2100
TGGTGGTTCG	GTCAACGGCG	GCGACGGCAC	CGACGTTCTG	GTGGCCAACG	TCAACGGTTC	2160
GTCGTTCAGC	GCTGACCCGG	CCTTCGGCGG	CTTCGAAACC	CTCCGCGTCG	CTGGCGCGGC	2220
GGCTCAAGGC	TCGCACAACG	CCAACGGCTT	CACGGCTCTG	CAACTGGCG	CGACGGCGGG	2280
TGCGACGACC	TTCACCAACG	TTGCGGTGAA	TGTGGCCTG	ACCGTTCTGG	CGGCTCCGAC	2340
CGGTACGACG	ACCGTACCCC	TGGCCAACGC	CACGGGCACC	TCGGACGTGT	TCAACCTGAC	2400
CCTGTCGTCC	TCGGCCGCTC	TGGCCGCTGG	TACGGTTGCG	CTGGCTGGCG	TCGAGACGGT	2460
GAACATCGCC	GCCACCGACA	CCAACACGAC	CGCTCACGTC	GACACGCTGA	CGCTGCAAGC	2520
CACCTCGGCC	AAGTCGATCG	TGGTGACGGG	CAACGCCGGT	CTGAACCTGA	CCAACACCGG	2580
CAACACGGCT	GTCACCAGCT	TCGACGCCAG	CGCCGTCACC	GGCACGGCTC	CGGCTGTGAC	2640
CTTCGTGTCC	GCCAACACCA	CGGTGGGTGA	AGTCGTACCG	ATCCGCGCG	GCGCTGGCGC	2700
CGACTCGCTG	ACCGGTTCGG	CCACCGCCAA	TGACACCATC	ATCGGTGGCG	CTGGCGCTGA	2760
CACCTGGTC	TACACCGCG	GTACGGACAC	CTTCACGGGT	GGCACGGCG	CGGATATCTT	2820
CGATATCAAC	GCTATCGGCA	CCTCGACCGC	TTTCGTGACG	ATCACCGACG	CCGCTGTGCG	2880
CGACAAGCTC	GACCTCGTCG	GCATCTCGAC	GAACGGCGCT	ATCGCTGACG	GCGCCTTCGG	2940
CGCTGCGGTC	ACCCCTGGCG	CTGCTGCGAC	CCTGGCTCAG	TACCTGGACG	CTGCTGCTGC	3000
CGGCGACGGC	AGCGGCACCT	CGGTTGCCAA	GTGGTTCCAG	TTCGGCGCG	ACACCTATGT	3060
CGTCGTTGAC	AGCTCGGCTG	GCGCGACCTT	CGTCAGCGGC	GCTGACGCGG	TGATCAAGCT	3120
GACCGGTCTG	GTCACGCTGA	CCACCTCGGC	CTTCGCCACC	GAAGTCCTGA	CGCTCGCCTA	3180
AGCGAACGTC	TGATCCTCGC	CTAGGCGAGG	ATCGCTAGAC	TAAGAGACCC	CGTCTTCCGA	3240
AAGGGAGGCG	GGGTCTTCT	TATGGCGCT	ACGCGCTGGC	CGGCCTTGCC	TAGTTCCGGT	3300

## Appendix 1 (cont'd)

Met Ala Tyr Thr Thr Ala Gln Leu Val Thr Ala Tyr Thr Asn Ala Asn  
 1 5 10 15  
 Leu Gly Lys Ala Pro Asp Ala Ala Thr Thr Leu Thr Leu Asp Ala Tyr  
 20 25 30  
 Ala Thr Gln Thr Gln Thr Gly Gly Leu Ser Asp Ala Ala Leu Thr  
 35 40 45  
 Asn Thr Leu Lys Leu Val Asn Ser Thr Thr Ala Val Ala Ile Gln Thr  
 50 55 60  
 Tyr Gln Phe Phe Thr Gly Val Ala Pro Ser Ala Ala Gly Leu Asp Phe  
 65 70 75 80  
 Leu Val Asp Ser Thr Thr Asn Thr Asn Asp Leu Asn Asp Ala Tyr Tyr  
 85 90 95  
 Ser Lys Phe Ala Gln Glu Asn Arg Phe Ile Asn Phe Ser Ile Asn Leu  
 100 105 110  
 Ala Thr Gly Ala Gly Ala Gly Ala Thr Ala Phe Ala Ala Ala Tyr Thr  
 115 120 125  
 Gly Val Ser Tyr Ala Gln Thr Val Ala Thr Ala Tyr Asp Lys Ile Ile  
 130 135 140  
 Gly Asn Ala Val Ala Thr Ala Ala Gly Val Asp Val Ala Ala Val  
 145 150 155 160  
 Ala Phe Leu Ser Arg Gln Ala Asn Ile Asp Tyr Leu Thr Ala Phe Val  
 165 170 175  
 Arg Ala Asn Thr Pro Phe Thr Ala Ala Asp Ile Asp Leu Ala Val  
 180 185 190  
 Lys Ala Ala Leu Ile Gly Thr Ile Leu Asn Ala Ala Thr Val Ser Gly  
 195 200 205  
 Ile Gly Gly Tyr Ala Thr Ala Thr Ala Ala Met Ile Asn Asp Leu Ser  
 210 215 220  
 Asp Gly Ala Leu Ser Thr Asp Asn Ala Ala Gly Val Asn Leu Phe Thr  
 225 230 235 240  
 Ala Tyr Pro Ser Ser Gly Val Ser Gly Ser Thr Leu Ser Leu Thr Thr  
 245 250 255  
 Gly Thr Asp Thr Leu Thr Gly Thr Ala Asn Asn Asp Thr Phe Val Ala  
 260 265 270  
 Gly Glu Val Ala Gly Ala Ala Thr Leu Thr Val Gly Asp Thr Leu Ser  
 275 280 285  
 Gly Gly Ala Gly Thr Asp Val Leu Asn Trp Val Gln Ala Ala Ala Val  
 290 295 300  
 Thr Ala Leu Pro Thr Gly Val Thr Ile Ser Gly Ile Glu Thr Met Asn  
 305 310 315 320  
 Val Thr Ser Gly Ala Ala Ile Thr Leu Asn Thr Ser Ser Gly Val Thr  
 325 330 335  
 Gly Leu Thr Ala Leu Asn Thr Asn Thr Ser Gly Ala Ala Gln Thr Val  
 340 345 350

## Appendix 1 (cont'd)

Thr Ala Gly Ala Gly Gln Asn Leu Thr Ala Thr Thr Ala Ala Gln Ala  
 355 360 365  
 Ala Asn Asn Val Ala Val Asp Gly Arg Ala Asn Val Thr Val Ala Ser  
 370 375 380  
 Thr Gly Val Thr Ser Gly Thr Thr Val Gly Ala Asn Ser Ala Ala  
 385 390 395 400  
 Ser Gly Thr Val Ser Val Ala Asn Ser Ser Thr Thr Thr Thr  
 405 410 415  
 Gly Ala Ile Ala Val Thr Gly Gly Thr Ala Val Thr Val Ala Gln Thr  
 420 425 430  
 Ala Gly Asn Ala Val Asn Thr Thr Leu Thr Gln Ala Asp Val Thr Val  
 435 440 445  
 Thr Gly Asn Ser Ser Thr Thr Ala Val Thr Val Thr Gln Thr Ala Ala  
 450 455 460  
 Ala Thr Ala Gly Ala Thr Val Ala Gly Arg Val Asn Gly Ala Val Thr  
 465 470 475 480  
 Ile Thr Asp Ser Ala Ala Ala Ser Ala Thr Thr Ala Gly Lys Ile Ala  
 485 490 495  
 Thr Val Thr Leu Gly Ser Phe Gly Ala Ala Thr Ile Asp Ser Ser Ala  
 500 505 510  
 Leu Thr Thr Val Asn Leu Ser Gly Thr Gly Thr Ser Leu Gly Ile Gly  
 515 520 525  
 Arg Gly Ala Leu Thr Ala Thr Pro Thr Ala Asn Thr Leu Thr Leu Asn  
 530 535 540  
 Val Asn Gly Leu Thr Thr Gly Ala Ile Thr Asp Ser Glu Ala Ala  
 545 550 555 560  
 Ala Asp Asp Gly Phe Thr Thr Ile Asn Ile Ala Gly Ser Thr Ala Ser  
 565 570 575  
 Ser Thr Ile Ala Ser Leu Val Ala Ala Asp Ala Thr Thr Leu Asn Ile  
 580 585 590  
 Ser Gly Asp Ala Arg Val Thr Ile Thr Ser His Thr Ala Ala Ala Leu  
 595 600 605  
 Thr Gly Ile Thr Val Thr Asn Ser Val Gly Ala Thr Leu Gly Ala Glu  
 610 615 620  
 Leu Ala Thr Gly Leu Val Phe Thr Gly Gly Ala Gly Arg Asp Ser Ile  
 625 630 635 640  
 Leu Leu Gly Ala Thr Thr Lys Ala Ile Val Met Gly Ala Gly Asp Asp  
 645 650 655  
 Thr Val Thr Val Ser Ser Ala Thr Leu Gly Ala Gly Ser Val Asn  
 660 665 670  
 Gly Gly Asp Gly Thr Asp Val Leu Val Ala Asn Val Asn Gly Ser Ser  
 675 680 685  
 Phe Ser Ala Asp Pro Ala Phe Gly Gly Phe Glu Thr Leu Arg Val Ala  
 690 695 700

23  
Appendix 1 (cont'd)

Gly Ala Ala Ala Gln Gly Ser His Asn Ala Asn Gly Phe Thr Ala Leu  
 705 710 715 720

Gln Leu Gly Ala Thr Ala Gly Ala Thr Thr Phe Thr Asn Val Ala Val  
 725 730 735

Asn Val Gly Leu Thr Val Leu Ala Ala Pro Thr Gly Thr Thr Val  
 740 745 750

Thr Leu Ala Asn Ala Thr Gly Thr Ser Asp Val Phe Asn Leu Thr Leu  
 755 760 765

Ser Ser Ser Ala Ala Leu Ala Ala Gly Thr Val Ala Leu Ala Gly Val  
 770 775 780

Glu Thr Val Asn Ile Ala Ala Thr Asp Thr Asn Thr Thr Ala His Val  
 785 790 795 800

Asp Thr Leu Thr Leu Gln Ala Thr Ser Ala Lys Ser Ile Val Val Thr  
 805 810 815

Gly Asn Ala Gly Leu Asn Leu Thr Asn Thr Gly Asn Thr Ala Val Thr  
 820 825 830

Ser Phe Asp Ala Ser Ala Val Thr Gly Thr Ala Pro Ala Val Thr Phe  
 835 840 845

Val Ser Ala Asn Thr Thr Val Gly Glu Val Val Thr Ile Arg Gly Gly  
 850 855 860

Ala Gly Ala Asp Ser Leu Thr Gly Ser Ala Thr Ala Asn Asp Thr Ile  
 865 870 875 880

Ile Gly Gly Ala Gly Ala Asp Thr Leu Val Tyr Thr Gly Gly Thr Asp  
 885 890 895

Thr Phe Thr Gly Gly Thr Gly Ala Asp Ile Phe Asp Ile Asn Ala Ile  
 900 905 910

Gly Thr Ser Thr Ala Phe Val Thr Ile Thr Asp Ala Ala Val Gly Asp  
 915 920 925

Lys Leu Asp Leu Val Gly Ile Ser Thr Asn Gly Ala Ile Ala Asp Gly  
 930 935 940

Ala Phe Gly Ala Ala Val Thr Leu Gly Ala Ala Ala Thr Leu Ala Gln  
 945 950 955 960

Tyr Leu Asp Ala Ala Ala Gly Asp Gly Ser Gly Thr Ser Val Ala  
 965 970 975

Lys Trp Phe Gln Phe Gly Gly Asp Thr Tyr Val Val Val Asp Ser Ser  
 980 985 990

Ala Gly Ala Thr Phe Val Ser Gly Ala Asp Ala Val Ile Lys Leu Thr  
 995 1000 1005

Gly Leu Val Thr Leu Thr Thr Ser Ala Phe Ala Thr Glu Val Leu Thr  
 1010 1015 1020

Leu Ala  
 1025

24  
Appendix 2

GAA TTC AGA TCT CAG GCC GCG GGG CAG GGT GGC TAT GGT GGG CTC GGC  
TCG CAA GGC  
GCT  
E F R S Q G A G Q G G Y G G L G S Q G A  
GGC CTG GGT GGC CAG GGC GCT GGC GCG GCC GCG GCC GCT GCG GCC GGT  
GGC  
G R G G Q G A G A A A A A A A A G G  
GCT GGC CAG GGC GGG CTG GGC TCG CAG GGC GCC GGC CAA GGC GCT GGC  
GCC GCG GCC  
GCT  
A G Q G G L G S Q G A G Q G A G A A A A  
GCG GCC GGT GGC GCC GGC CAG GGT GGC TAC GGC GGC CTG GGC AGC CAG  
GGC GCC GGT  
CGC  
A A G G A G Q G G Y G G L G S Q G A G R  
GGC GGT CAG GGC GCC GGT GGC GCG GCC GCT GCG GCC GGT GGC GCT GGG  
CAA GGC GGC TAC  
G G Q G A G A A A A A G G A G Q G G Y  
GGC GGT CTG GGA TCC  
G G L G S

1/1

## Appendix 3

atg aac aca aac aag gca acc gca act tac ttg aaa tcc att atg ctt cca gag act  
gga  
Met asn thr asn lys ala thr ala thr tyr leu lys ser ile met leu pro glu thr  
gly  
61/21

cca gca agc atc ccg gac gac ata acg gag aga cac aic tta aaa caa gag acc tcc  
tca  
pro ala ser ile pro asp asp ile thr glu arg his ile leu lys gln glu thr ser  
ser  
121/41

tac aac tta gag gtc tcc gaa tca gga agt ggc att ctt gtt tgc cct ggg gca  
cca  
tyr asn leu glu val ser glu ser gly ser gly ile leu val cys phe pro gly ala  
pro  
181/61

ggc tca cgg atc ggt gca cac tac aga tgg aat grg aac cag acg ggg ctg gag ttc  
gac  
gly ser arg ile gly ala his tyr arg trp asn ala asn gln thr gly leu glu phe  
asp  
241/81

cag tgg ctg gag acg tcc cag gac ctg aag aaa gcc ttc aac tac ggg agg ctg atc  
tca  
gln trp leu glu thr ser gln asp leu lys lys ala phe asn tyr gly arg leu ile  
ser  
301/101

agg aaa tac gac att caa agc tcc aca cta ccg gcc ggt ctc tat gct ctg aac ggg  
acg  
arg lys tyr asp ile gln ser ser thr leu pro ala gly leu tyr ala leu asn gly  
thr  
361/121

ctc aac gct gcc acc ttc gaa ggc agt ctg tct gag gtg gag agc ctg acc tac aat  
acg  
leu asn ala ala thr phe glu gly ser leu ser glu val glu ser leu thr tyr asn  
ser  
421/141

ctg atg tcc cta act acg aac ccc cag gac aaa gcc aac aac cag ctg gtg acc aaa  
gga  
ieu met ser leu thr thr asn pro gln asp lys ala asn asn gln leu val thr lys  
gly  
481/161

gtc acc gtc ctg aat cta cca aca ggg ttc gac aaa cca tac gtc cgc cta gag gac  
gag  
val thr val leu asn leu pro thr gly phe asp lys pro tyr val arg leu glu asp  
glu  
541/181

aca ccc cag ggt ctc cag tca atg aac ggg gcc agg atg agg tgc aca gct gca att  
gca  
thr pro gln gly leu gln ser met asn gly ala arg met arg cys thr ala ala ile  
ala  
601/201

cca cgg agg tac gag atc gac ctc cca caa agc cta ccc ccc gtt cct gcg aca  
gga  
pro arg arg tyr glu ile asp leu pro ser gln ser leu pro pro val pro ala thr  
gly  
661/221

acc ctc acc act ctc tac gag gga aac gcc gac atc gtc agc tcc aca aca gtg acg  
gga  
thr leu thr leu tyr glu gly asn ala asp ile val ser ser thr thr val thr  
gly  
721/241

gac ata aac ttc agt ctg gca gaa cga ccc gca aac gag acc agg ttc gac ttc cag  
ctg  
asp ile asn phe ser leu ala glu arg pro ala asn glu thr arg phe asp phe gln  
leu

26  
Appendix 4

The T3 protein sequence is:

FACKTANGTAIPIGGGSANVYVNLAPVNVGQNLVVDLSTQIFCHNDYPETITDYVTLQRGSA  
SYPFPTTSETPRVYNSRTDKPWPVALYLTPVSSAGGVAIKAGSLIAVLILRQTNNYNSDDFQ  
CDVSA

The T7 protein sequence is:

FACKTANGTAIPIGGGSANVYVNLAPVNVGQNLVVDLSTQIFCHNDYPETITDYVTLQRGSA  
SYPFPTTSETPRVYNSRTDKPWPVALYLTPVSSAGGVAIKAGSLIAVLILRQTNNYNSDDFQ  
CDVSARDVTVPDYRGSPVPIPLTVYCAKSQNLGYYLSGTHADAGNSIFTNTASFSPAQGVG  
GAVGTSAVSLGLTANYARTGGQVTAGNVQSIIGVTFVYQ

**WHAT IS CLAIMED IS:**

1. A method of cleaving a fusion protein including a first component which comprises all or part of a Caulobacter S-layer protein including a Caulobacter C-terminal secretion signal, and a second component heterologous to Caulobacter, the fusion protein containing at least one aspartate-proline dipeptide, wherein the method comprises combining the fusion protein with an acid solution of a strength insufficient to solubilize the fusion protein for a time sufficient for cleavage of the fusion protein at said aspartate-proline dipeptide.
- 10 2. The method of claim 1 wherein a aspartate-proline dipeptide is situated between the first and second components or adjacent a junction between the first and second components.
- 15 3. The method of claim 1 or 2, wherein the acid solution has a pH of from about 1.5 to about 2.5.
4. The method of claim 1 or 2, wherein the acid solution has a pH of about 1.65 to about 2.35.
- 20 5. The method of any one of claims 1-4 wherein the method is carried out at a temperature in the range of about 30° C. to about 50° C.
6. The method of any one of claims 1-5, wherein the method further comprises separating products cleaved from the fusion protein.
- 25 7. A method of preparing a DNA construct for expression of a fusion protein suitable for use in the method of claim 1, wherein the method comprises joining an upstream DNA segment including DNA heterologous to Caulobacter which encodes a protein

of interest, to a downstream DNA segment including DNA for a Caulobacter C-terminal secretion signal, wherein the downstream DNA segment does not encode an aspartate-proline dipeptide, and wherein the upstream segment contains DNA encoding an aspartate-proline dipeptide at or near an end of said upstream segment to 5 be joined to said downstream segment.

8. A method of preparing a fusion protein, comprising:

(1) expressing a DNA construct prepared as described in claim 7 in Caulobacter and,

10

(2) recovering said fusion protein secreted by the Caulobacter.

# INTERNATIONAL SEARCH REPORT

Inter. Application No  
PCT/CA 99/00637

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/62 C12N15/74 C07K14/21 C07K14/00 C07K14/08  
C07K14/245

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
IPC 7 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the International search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 97 34000 A (UNIV BRITISH COLUMBIA ;SMIT JOHN (CA); BINGLE WADE H (CA); NOMELLI) 18 September 1997 (1997-09-18) cited in the application the whole document	1-8
Y	US 5 500 353 A (SMIT JOHN ET AL) 19 March 1996 (1996-03-19) the whole document	1-8 -/-

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

### \* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the International filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the International filing date but later than the priority date claimed

- "T" later document published after the International filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "Z" document member of the same patent family

Date of the actual completion of the International search

17 December 1999

Date of mailing of the International search report

11/01/2000

Name and mailing address of the ISA  
European Patent Office, P.B. 5018 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax. (+31-70) 340-3016

Authorized officer

Chambonnet, F

## INTERNATIONAL SEARCH REPORT

Int'l. Search Application No  
PCT/CA 99/00637

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	<p>BINGLE W H ET AL: "LINKER MUTAGENESIS OF THE CAULOBACTER CRESCENTUS S-LAYER PROTEIN: TOWARD A DEFINITION OF AN N-TERMINAL ANCHORING REGION AND A C-TERMINAL SECRETION SIGNAL AND THE POTENTIAL FOR HETEROLOGOUS PROTEIN SECRETION"</p> <p>JOURNAL OF BACTERIOLOGY, US, WASHINGTON, DC, vol. 179, no. 3, February 1997 (1997-02), page 601-611 XP002034945</p> <p>ISSN: 0021-9193</p> <p>cited in the application</p> <p>page 604, column 1, line 5 -page 608, column 1, paragraph 2; figure 2</p>	1-8
Y	<p>NOMELLINI J F ET AL: "INSERTION OF HETEROLOGOUS PEPTIDES WITHIN THE SURFACE-LAYER PROTEIN OF CAULOBACTER CRESCENTUS"</p> <p>ABSTRACTS OF THE GENERAL MEETING OF THE AMERICAN SOCIETY FOR MICROBIOLOGY, US, THE SOCIETY, WASHINGTON, DC, vol. 95, 1995, page 525 XP002035105</p> <p>ISSN: 1060-2011</p> <p>abstract</p>	1-8
Y	<p>JOHN WILEY &amp; SONS INC.: "Introduction to expression by fusion protein vectors"</p> <p>CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, vol. suppl 28, no. chapter 16.4A, pages 16.4.1-16.4.17, XP002125707</p> <p>cited in the application</p> <p>page 16.4.1 -page 16.4.4</p> <p>page 16.4.13</p>	1-8
Y	<p>LANDON M. : "Cleavage at Aspartyl-Prolyl bonds"</p> <p>METHODS IN ENZYMOLOGY, vol. XLVII, no. part E, 1977, pages 145-149, XP000863051</p> <p>cited in the application</p> <p>the whole document</p>	1-8

## INTERNATIONAL SEARCH REPORT

Information on patent family members

Int'l. Application No.

PCT/CA 99/00637

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO 9734000	A	18-09-1997	US 5976864 A	02-11-1999
			US 5500353 A	19-03-1996
			AU 2019497 A	01-10-1997
			CA 2175549 A	13-09-1997
			CA 2247905 A	19-09-1997
			EP 0888454 A	07-01-1999
			CA 2090549 A	10-12-1999
US 5500353	A	19-03-1996	WO 9734000 A	18-09-1997
			US 5976864 A	02-11-1999
			CA 2090549 A	10-12-1993